

# UC Irvine

## UC Irvine Previously Published Works

### Title

Character space restrictions and boundary conditions in the evolution of quantitative multistate characters.

### Permalink

<https://escholarship.org/uc/item/1ht648x6>

### Journal

Journal of theoretical biology, 80(1)

### ISSN

0022-5193

### Authors

Chapman, RW  
Avisé, JC  
Asmussen, MA

### Publication Date

1979-09-01

### DOI

10.1016/0022-5193(79)90179-6

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

## Character Space Restrictions and Boundary Conditions in the Evolution of Quantitative Multistate Characters

ROBERT W. CHAPMAN, JOHN C. AVISE AND  
MARJORIE A. ASMUSSEN

*Department of Zoology, University of Georgia,  
Athens, Georgia 30602, U.S.A.*

*(Received 31 January 1979)*

The effects of restrictions of available character space on the mean morphological distance between living members of evolutionary phylads are examined by Monte Carlo simulation. The approach involves specifying the degree to which ancestor-descendant species may differ and limiting the range of attainable character states within a phylad. Morphological evolution is modeled as a Markovian process involving quantitative multistate characters. States for a given character are allowed to evolve at time-dependent or speciation-dependent rates.

The final distributions of morphological distance for a given trait among members of a phylad depend on the number of species in the phylad, the rate and pattern of evolution of new character states, and the existence of boundary conditions indicating possible selective constraints on the trait. When morphological change is proportional to time, increasing restrictions on character evolution tend to (a) lower mean distance between species and (b) leave the ratio of mean distances ( $\bar{D}_R/\bar{D}_P$ ) in species-rich vs. species-poor phylads of comparable evolutionary age near one. When change is proportional to rate of speciation, similar restrictions tend to (a) limit mean distance only in phylads in which the number of speciations exceeds the range of attainable character states and (b) permit  $\bar{D}_R/\bar{D}_P$  to be considerably greater than one, except in extreme cases. Implications of these results for the current phyletic gradualism-rectangular evolution controversy are considered.

### 1. Introduction

Avise & Ayala (1975) introduced theoretical models which led to predictions about whether evolutionary divergence is a function of time or of speciational events. The models consider pairs of phylads of equal age, one species-rich and the other species-poor. If divergence is proportional to rate of speciation, then random pairs of living species from a species-rich phylad

should be more distinct than random pairs of living species from a species-poor phylad, since on the average, they will be separated by more speciations. Alternatively, if divergence is proportional to time since species last shared a common ancestor, pairs of living species in both phylads should be separated by about the same distance, on the average. In general, if divergence is a function of time, the ratio of average distances ( $\bar{D}_R/\bar{D}_P$ ) in species-rich ( $\bar{D}_R$ ) and species-poor ( $\bar{D}_P$ ) phylads of similar evolutionary age should be near unity, but if divergence is a function of speciation events, the ratio  $\bar{D}_R/\bar{D}_P$  will be  $> 1$  (Avice & Ayala, 1975).

One criticism of this approach is directed to an assumption of the models that no convergence occurs. During the evolutionary process, characters (morphological, genetic, etc.) change from one state to another, but these changes will be reflected in an increase in distance between species only in the unlikely event that convergence is non-existent. The severity of convergence depends on a variety of factors which can be modeled in an attempt to increase the reality of the models' predictions. Avice (submitted) relaxed the "no convergence" assumption for qualitative multistate characters by using a deterministic approach which limits to various degrees the number of character states attainable by members of a phylad. The results showed that  $\bar{D}_R/\bar{D}_P$  remains near unity when differentiation is a function of time, even when convergence is severe. The ratio with character convergence also approaches unity when differentiation is a function of speciational events, but the effect is not severe unless the number of character states is very small with respect to size of the phylad, and the probability per speciation of a change from one character state to another is very large.

This paper extends the study of character convergence to quantitative multistate traits. At the outset a distinction must be made between qualitative and quantitative characters. By qualitative characters, we mean those which cannot be arrayed in some obvious order and hence are only characterized as alike or different (see Sneath & Sokal, 1973). Electromorph products of specific structural gene loci are characters of this type. We define quantitative multistate traits as characters expressed by a numerical value which can be arranged in order of magnitude along a one-dimensional axis. Quantifiable morphological traits, both continuous and meristic, are often arranged in such ordered sets. The evolution of quantitative traits may be visualized and modeled as a path dependent (Markovian) process (Raup *et al.*, 1973; Raup & Gould, 1974; Raup, 1977). We will use Monte Carlo simulations to determine if the approach of Avice & Ayala (1975) will generate distinct predictions about mean morphological distance within a phylad under the competing evolutionary scenarios of (1) time and (2) speciation dependent rates of morphological divergence.

## 2. The Models

Particularly when the number or range of attainable states for a given character is much smaller than the number of living species in a phylad, stringent limits must exist on the level of phenetic divergence among species for that trait. All else being equal, as the ratio  $N_S/N_C$  increases, where  $N_C$  is the number of attainable character states and  $N_S$  is the number of living species, the absolute distance between many pairs of species must decrease. The effect of this restriction on morphological distance within a phylad depends not only on the number of attainable character states and the number of species, but also on the rate and pattern by which the character states are arrived at in the evolutionary process.

We consider a fixed branching process (Fig. 1) where branching (speciation) is dichotomous and occurs at regular time intervals. This branching pattern is a special case of more general patterns (cf. Gould *et al.*, 1977) where the center of gravity (CG, a measure of mean diversity on a time axis 0, 1) is  $> 0.5$  and uniformity (UNI, a measure of fluctuation in diversity) is  $< 0.5$ . Both UNI and CG are more fully explained in Gould *et al.* (1977). In addition, we assume that no lineages go extinct and character states change each time interval or speciation event, depending on the model.

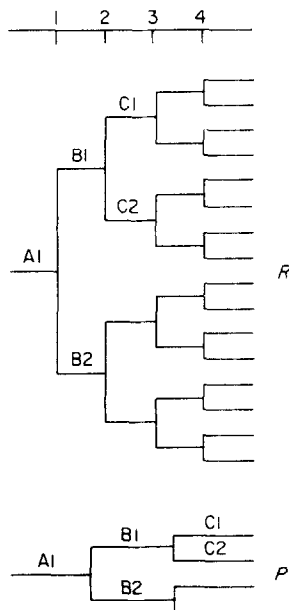


FIG. 1. Two phylads, one species-rich (*R*) and the other species-poor (*P*). Cladogenetic events give rise to two new species and occur each time interval in *R* and every 1.75 time intervals in *P*.

## (A) DIVERGENCE AS A FUNCTION OF SPECIATION (THE CLAD MODEL)

A simplified flow diagram for the computer simulation of this model is shown in Fig. 2. The program assigned a distance of one unit between ancestor and descendant species separated by one speciation. The probability of increase or decrease from the ancestral character state was equal to 0.5; hence  $B_1$  and  $B_2$  (Fig. 1) may share a state (either  $A_1 + 1$  or  $A_1 - 1$ ), or differ by two units  $[(A_1 + 1) - (A_1 - 1)]$ , with equal likelihood. The outcome was determined by automatically generated random numbers. The

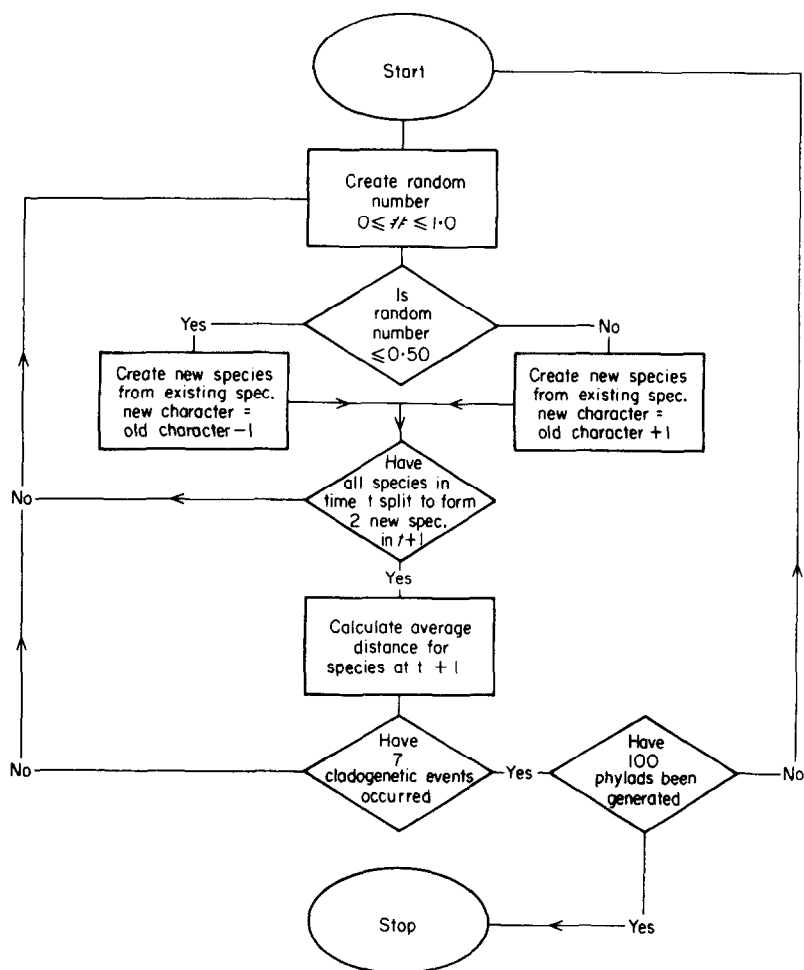


FIG. 2. Simplified flow diagram for Monte Carlo simulation of the clad model.

TABLE 1

*Mean average distance ( $\bar{D}$ ) from 100 runs of the Monte Carlo simulations for time and clad models (see text)*

No. species	Clad model			Time model		
	Unbounded	$\pm 5$	$\pm 3$	Unbounded	$\pm 5$	$\pm 3$
4	1.276	1.437	1.183	2.213	2.119	1.882
8	1.647	1.684	1.544	2.330	2.255	1.830
16	1.937	2.000	1.638	2.595	2.455	1.917
32	2.228	2.263	1.879	2.477	2.387	1.920
64	2.469	2.453	1.948	2.475	2.476	1.921
128	2.712	2.649	2.084	2.712	2.649	2.084

simulation continued for seven speciation (cladogenetic) episodes where 128 species were extant after the final branching.

In principle, character spaces of two types are available to evolutionary phylads: those that have yet to be saturated by the extant species of the phylad and those that can be saturated with the existing numbers of species and rates of evolutionary change. We considered two forms of the branching process which simulate these types of characters. First, we allowed the range of attainable character states to be limited only by the number of cladogenetic events. Hence, after seven cladogenetic episodes, the range of attainable character states was  $\pm 7$  from the origin. Second, we constrained the range of character states to boundaries of  $\pm 3$  or  $\pm 5$  from the origin. At the boundaries, transition probabilities were altered to allow descendants to maintain the boundary character state or to "reflect" inward, again with equal probability. Under these constraints only splitting patterns in which more than three or five cladogenetic events occurred were affected.

TABLE 2

*Distance ratios ( $\hat{D}_R/\hat{D}_P$  and, in parentheses,  $\bar{D}_R/\bar{D}_P$ ) from Monte Carlo simulations for both time and clad models (see text)*

$N_{SR}/N_{SP}$	Unbounded		$\pm 5$		$\pm 3$	
	Clad	Time	Clad	Time	Clad	Time
128/8	1.779 (1.647)	1.289 (1.164)	1.840 (1.573)	1.259 (1.175)	1.548 (1.350)	1.374 (1.139)
128/16	1.414 (1.400)	1.130 (1.045)	1.429 (1.325)	1.303 (1.079)	1.380 (1.272)	1.214 (1.087)
128/32	1.245 (1.217)	1.128 (1.094)	1.220 (1.171)	1.135 (1.109)	1.162 (1.109)	1.130 (1.085)
128/64	1.133 (1.098)	1.231 (1.096)	1.109 (1.079)	1.261 (1.069)	1.103 (1.069)	1.183 (1.085)

For each phylad and set of conditions, an average distance ( $\bar{d}$ ) between extant species was calculated by an Euclidian distance metric ( $\Delta_{jk}$ , p. 124, Sneath & Sokal, 1973). The program was repeated 100 times, generating mean average distances ( $\bar{D} = \sum d/100$ ) for phylads consisting of 4, 8, 16, 32, 64 and 128 species (Table 1). The ratios of mean distances in species-rich vs. species-poor phylads of comparable evolutionary age ( $\bar{D}_R/\bar{D}_P$ ) are presented in Table 2.

In an effort to conserve core storage (i.e. cost), 70 randomly chosen average distances for phylads containing 8 to 64 species were compared to the same number of distances for the 128 species phylads. This yielded, for each pairwise comparison of phylads, 4900 ratios of  $\bar{d}_R/\bar{d}_P$  and permitted another estimate of mean ratios of distance in species-rich vs. species-poor phylads:

$$\hat{D}_R/\hat{D}_P = \frac{\sum \bar{d}_R/\bar{d}_P}{4900}.$$

(The 128/2 and 128/4 ratios were not calculated since many of the denominator distances were 0.) In general, for any pair of phylads, the ratios  $\bar{D}_R/\bar{D}_P$  and  $\hat{D}_R/\hat{D}_P$  are similar (Table 2).

Simulation programs are difficult to debug, as there are usually no exact answers to check against the results. However, in this simulation an "exact" answer for the simple cases of 2 and 4 species phylads can be obtained by an appeal to probability theory. Using the formula

$$E(\bar{D}) = \sum_{i=1}^n p_i \bar{d}_i,$$

where  $p_i$  is the probability of the "ith" outcome,  $\bar{d}_i$  is the average distance for "i" and  $n$  is the number of possible outcomes, an expected average distance  $E(\bar{D})$  can be obtained. The computer simulation results were close to expectations for the 2-species ( $E(\bar{D}) = 1.00$ , simulated  $\bar{D} = 0.96$ ) and 4-species phylads ( $E(\bar{D}) = 1.33$ , simulated  $\bar{D} = 1.28$ ). Remaining results were accepted on these grounds.

#### (B) DIVERGENCE AS A FUNCTION OF TIME (THE TIME MODEL)

The computer simulation for this model was somewhat more complex to program. The simulation allows a taxon to diverge one unit from its ancestral state per time interval. At branching points (Fig. 1), the sister-taxa diverged one unit from the ancestral state per time interval. Thus after one time interval the distance between sister species could be 0, 2; after two time intervals 0, 2, 4, etc. To remain consistent with the clad divergence model,

splitting patterns generating 2–128 species were allowed to evolve through seven time intervals. Speciation events occurred at regular intervals and average distances were calculated at the end of the run. The data for this model were summarized as before (Tables 1 and 2). The simulation was checked against an analytical solution for the two species phylad, the only branching pattern for which a solution was feasible ( $E(\bar{D}) = 2.18$ , simulated  $\bar{D} = 2.42$ ).

As in the previous model, we considered the same criteria for boundary conditions. Since all branching patterns were capable of attaining the full range of character states, larger distance values were unlikely to occur in all branching patterns. Thus, the limitation in character states produced by boundary conditions affected all branching patterns, not just those in which more than three or five cladogenetic events occurred.

Since it is theoretically possible for  $\bar{d}_R/\bar{d}_P$  to assume a broad range of values in both the clad and time models, we wish to know the likelihood a given ratio can be assigned to either model. Frequency distributions of  $\bar{d}_R/\bar{d}_P$  (in 0.1 unit increments) were recorded (Table 3) and plotted (Figs 3 and 4). To obtain a likelihood with which a given ratio can be ascribed to the time model, the number of observed outcomes with a given ratio in the time model can be divided by the sum of outcomes with that ratio in both models.

Computer programs were written in FORTRAN and executed on an IBM-370 computer at the University of Georgia. Copies of flow charts and source decks may be obtained from one of us (RWC).

### 3. Results

The effect of path-dependent evolution on  $\bar{D}$ 's in quantitative multistate characters within an evolutionary phylad is strongly model dependent. When the number of attainable character states is a function of rate of speciation,  $\bar{D}$  increases as the number of speciation events increases (Table 1). In the time divergence model,  $\bar{D}$  remains nearly constant regardless of the number of speciation events. When moderate ( $\pm 5$ ) boundary conditions are imposed, average distances in both models are relatively unaffected. More stringent conditions ( $\pm 3$ ) significantly reduce  $\bar{D}$  in the time divergence model and in the clad divergence model when four or more speciation episodes occur (Table 1).

When divergence is a function of time,  $\bar{D}_R/\bar{D}_P$  (or  $\hat{D}_R/\hat{D}_P$ ) remains near one, irrespective of the number of species in the phylads compared or restrictions on range of attainable character states (Table 2). In the clad divergence model,  $\bar{D}_R/\bar{D}_P$  (or  $\hat{D}_R/\hat{D}_P$ ) approaches one as differences in the rate of





1-6	237	156	143	64	459	425	316	107	214	187	145	158	334	383	219	121	102	118	78	50	253	241	139	99
1-7	185	102	71	52	335	357	167	67	174	130	92	101	350	290	142	59	58	78	51	34	244	206	96	51
1-8	152	60	25	32	317	249	91	42	152	111	38	56	335	214	102	20	46	77	35	30	157	181	68	9
1-9	182	55	8	38	360	174	57	18	115	72	23	50	232	153	56	17	58	55	26	41	239	127	46	10
2-0	120	34	4	28	229	155	25	15	70	71	10	39	186	142	33	8	53	46	15	42	198	136	28	—
2-1	68	16	—	15	108	115	7	9	74	44	5	27	209	102	19	3	45	70	17	48	201	114	19	—
2-2	79	23	—	25	152	84	—	1	46	38	—	15	125	64	9	4	47	78	17	41	162	80	23	—
2-3	54	22	—	21	115	79	—	—	46	25	1	13	143	51	12	1	47	63	7	29	57	68	8	—
2-4	67	21	—	25	126	46	—	—	19	23	—	15	52	27	5	—	52	50	8	34	84	67	3	—
2-5	37	30	—	18	98	36	—	—	38	8	—	14	111	27	7	—	44	48	9	45	39	47	2	—
2-6	37	17	—	20	89	19	—	—	27	13	—	14	59	14	1	—	35	38	4	31	36	40	—	—
2-7	34	28	—	23	104	4	—	—	29	9	—	7	94	17	4	—	48	23	6	23	15	18	—	—
2-8	21	15	—	17	62	3	—	—	13	5	—	5	46	5	—	—	64	18	3	16	15	12	—	—
2-9	13	20	—	21	74	—	—	—	21	6	—	7	65	11	1	—	49	13	9	19	9	13	—	—
3-0	6	15	—	11	75	—	—	—	10	2	—	5	25	5	—	—	31	7	9	14	9	15	—	—
3-1	2	16	—	8	48	—	—	—	13	4	—	7	30	5	—	—	34	4	9	13	9	6	—	—
3-2	—	6	—	10	47	—	—	—	7	2	—	4	20	9	—	—	19	4	4	5	3	10	—	—
3-3	—	2	—	7	17	—	—	—	6	2	—	6	8	6	—	—	24	6	7	7	0	4	—	—
3-4	—	—	—	3	32	—	—	—	5	2	—	3	25	8	—	—	6	3	2	2	3	4	—	—
3-5	—	—	—	3	20	—	—	—	—	5	—	—	6	4	—	—	14	4	2	2	6	1	—	—
≥3-6	—	—	—	72	38	—	—	—	9	130	—	75	17	10	—	—	80	45	2	72	177	4	—	—

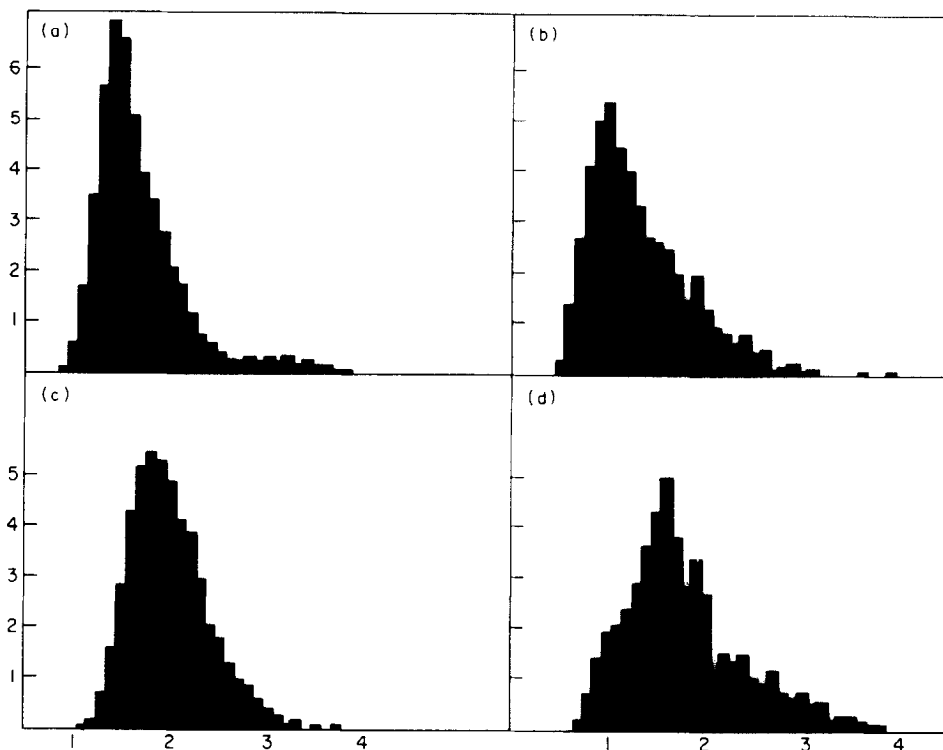


FIG. 3. Frequency distribution of distance ratios ( $\bar{d}_R/\bar{d}_P$ ) for unbounded time (a, b) and clad (c, d) models in 0.1 intervals. Vertical axis is in hundreds. a, c = 128-species vs. 16-species phylads; b, d = 128 vs. 8.

speciation in the phylads compared decreases, but restrictions on the range of character states ( $\pm 5$ ,  $\pm 3$ ) have little effect (Table 2).

From Table 3, it can be shown that  $\bar{d}_R/\bar{d}_P \leq 1.0$  occurred more often in the time model than clad model, especially when the phylads contained vastly different numbers of species. The probability that  $\bar{d}_R/\bar{d}_P \leq 1.0$  can be ascribed to the time model is  $\geq 0.72$  (unbounded),  $\geq 0.62$  (bounded  $\pm 5$ ), or  $\geq 0.59$  (bounded  $\pm 3$ ) when the ratio of number of species in the phylads ( $N_R/N_P$ ) was  $\geq 8$ . Discrimination between the time model and clad model is not clear (and on occasion reversed from the above) when  $N_R/N_P \leq 4$ .  $\bar{d}_R/\bar{d}_P \geq 1.4$  occurs more often in the clad model, but discrimination is not particularly good and diminishes as boundary conditions become more stringent or as  $N_R/N_P$  decreases.

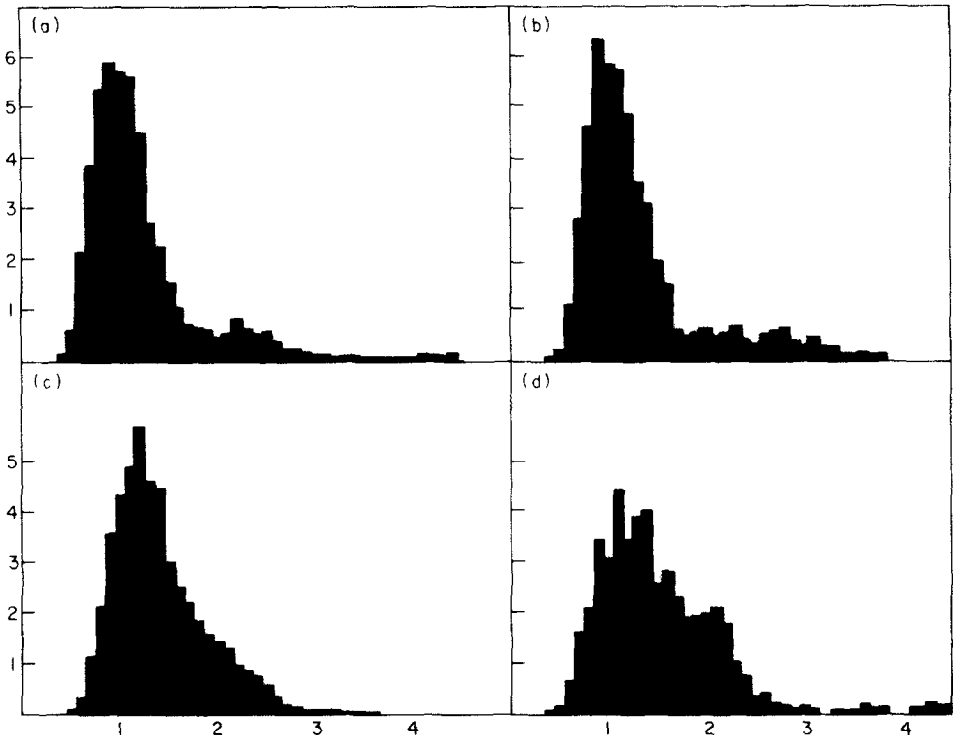


FIG. 4. Frequency distributions of distance ratios ( $\bar{d}_R/\bar{d}_P$ ) for time and clad models with boundary conditions  $\pm 3$ . See caption to Fig. 3.

#### 4. Discussion

A current controversy in evolutionary biology concerns whether most evolutionary change is gradual and occurs by the slow and even transformation of populations within established species (phyletic gradualism), or whether most evolutionary change is concentrated in speciation episodes (rectangular evolution, or evolution through punctuated equilibria) (Eldredge & Gould, 1972; Gould & Eldredge, 1977; Stanley, 1975). The proposal of rectangular evolution was forged largely out of a literal interpretation of the fossil record, with "gaps" interpreted to represent the logical consequences of rapid changes occurring in geographically localized populations undergoing speciation. Thus any empirical test of the rectangular evolution-phyletic gradualism controversy which rests wholly or in large part upon a literal interpretation of fossil remains would appear strongly biased in favor of the punctuated equilibrium model. It therefore seems desirable to develop additional hypotheses and tests, not entirely

dependent upon fossil evidence, to resolve which mode of evolution is responsible for the majority of evolutionary change.

As discussed in the introduction, the original models of Avise & Ayala (1975) provide an approach to the controversy which involves comparisons among living representatives of evolutionary phylads. Before predictions of these theoretical models can be tested with empirical data from evolutionary phylads which have speciated at different rates, it is imperative to determine the relationship of "increase in distance" on the one hand, to the "changes in character states" which actually do occur during the evolutionary process. The two are identical only when character convergence is non-existent.

In this paper, we use Monte Carlo computer simulations to examine the effects of restrictions on available character space in quantitative multistate traits, when these traits are allowed to change in time-dependent vs. speciation-dependent manner. Two general types of constraints prohibit "changes in character states" from being directly translatable into increased distance in Euclidian space: (a) any given quantitative character can only increase or decrease in numerical value along a one-dimensional axis which is nonetheless occupied by many species and (b) even along this axis, boundary conditions (due to natural selection) may exist which prohibit unrestricted increases or decreases in the value a trait can assume.

As a concrete example, consider the number of anal fin rays (or almost any other quantifiable continuous or meristic trait) in a phylad of fishes. Whether changes in fin ray number among these species of fish have occurred at time-dependent or speciation-dependent rates, only a very finite number of fin ray numbers can possibly yield a functional swimming appendage. For example, the North American cyprinid fishes comprise some 250 species, yet all exhibit between 6 and 12 anal fin rays (7–11 dorsal fin rays, 30–90 lateral line scales, etc.). Similar observations could be made for a variety of characters in this and other phylads.

Results of the computer simulations show that when changes in states of morphological characters are time-dependent processes, the ratio of mean Euclidian distance among living representatives of rapidly speciating vs. slowly speciating phylads of similar evolutionary age ( $\bar{D}_R/\bar{D}_P$ ) remains slightly greater than unity. On the contrary, when morphological change occurs at speciation-dependent rates,  $\bar{D}_R/\bar{D}_P$  may be considerably greater than one, particularly when the difference in rates of speciation of the two phylads under consideration is great, and when the boundary conditions are weak or non-existent. Although we have examined only a limited number of evolutionary phylads due to cost of the computer simulations, the results match our intuitive appraisal and offer some hope that rectangular and phyletic gradualism models may be empirically and critically evaluated by

examining morphological characteristics in appropriately chosen phylads. The distinctness of the models' predictions are not, however, as great as are those for qualitative multistate characters, or in situations in which character convergence is non-existent. In the case of quantitative characters, the overlap in frequency distribution of  $\bar{d}_R/\bar{d}_P$  may be considerable, and results of empirical tests must be interpreted with caution.

Another cause for reservations about empirical tests of these models concerns our scenario about the shape of evolutionary phylads,  $CG \geq 0.5$  and  $UNI \leq 0.5$ . Some real phylads may undergo explosive diversification early in their history, reach a saturation density (analogous to carrying capacity in population ecology), and maintain relatively stable numbers of species for considerable lengths of time (cf. Sepkoski, 1978; Stanley, 1978) leading to phylads with  $UNI \geq 0.5$  (Gould *et al.*, 1977). Judging only from numbers of extant species, it would be impossible to decide whether the different numbers of species in two phylads resulted from differing speciation rates or from different saturation densities. Yet this difference could be critical to predictions of  $\bar{D}_R/\bar{D}_P$ . In the clad model, if morphologies are restricted to certain bounds and if speciation balanced by extinction occurs after saturation density is achieved, character states will tend to be equitably distributed among species in each phylad.

A variety of yardsticks may be employed to measure evolutionary divergence. We have previously tested the rectangular mode of evolution with respect to products of structural genes in two families of North American fishes apparently characterized by different rates of speciation (Avice, 1977). Results suggested that rates of protein evolution in these fishes were compatible with phyletic gradualism and probably incompatible with punctuated equilibrium, but these conclusions do not necessarily hold for other aspects of phenotype or genotype. The real proposal of proponents of rectangular evolution is that most *morphological* change in evolution occurs during speciation episodes. Since most morphological characteristics are quantifiable as either continuous or meristic traits, the present models indicate that the rectangular mode of evolution may also be tested using morphological characters in living representatives of carefully chosen phylads.

This study was supported in part by a Genetics training grant to RWC, PHS No. 5-Tcw-GM07103-04. The assistance of Dr T. H. Emigh in various areas of the work is gratefully acknowledged.

## REFERENCES

- AVISE, J. C. (1977). *Proc. natn. Acad. Sci. U.S.A.* **74**, 5083.  
AVISE, J. C. & AYALA, F. J. (1975). *Genetics* **81**, 757.

- ELDRIDGE, N. & GOULD, S. J. (1972). In *Models of Paleobiology* (T. J. M. Schopf, ed.), pp. 82-115. San Francisco: Freeman, Cooper and Co.
- GOULD, S. J. & ELDRIDGE, N. (1977). *Paleobiology* **3**, 115.
- RAUP, D. M. (1977). *Am. Sci.* **65**, 50.
- RAUP, D. M., GOULD, S. J., SCHOPF, T. J. M. & SIMBERLOFF, D. S. (1973). *J. Geol.* **81**, 525.
- RAUP, D. M. & GOULD, S. J. (1974). *Syst. Zool.* **23**, 305.
- SEPKOSKI, J. J., Jr. (1978). *Paleobiology* **4**, 223.
- SNEATH, P. H. A. & SOKAL, R. R. (1973). *Numerical Taxonomy*, 573 pp. San Francisco: W. H. Freeman & Co.
- STANLEY, S. M. (1975). *Proc. natn. Acad. Sci. U.S.A.* **72**, 646.
- STANLEY, S. M. (1978). *Paleobiology* **4**, 26.